doi:10.6041/j.issn.1000-1298.2026.12.022

基于变分推理半监督学习的玉米雄穗实例分割方法

王奇瑞[1,2]　王　龙[2]　宋时芳[3]　周恩权[2,3]　毛罕平[3]　刘　洋[4]

(1.常州工学院机械工程学院,常州 213032;2.海安上海交通大学智能装备研究院,海安 226600;
3.江苏大学现代农业装备与技术教育部重点实验室,镇江 212013;4.新疆农垦科学院机械装备研究所,石河子 832000)

摘要:针对玉米雄穗实例分割任务中标注数据制作耗时费力、易受主观因素影响,以及田间复杂场景导致预测不确定性高的问题,本文提出一种自适应变分推理贝叶斯实例分割模型(AVB-IS)。AVB-IS 将贝叶斯概率框架与深度学习相融合,通过变分推理学习玉米雄穗特征分布,并通过自适应先验优化潜在空间,使模型在单次前向传播中即可预测不确定性。AVB-IS 利用不确定性信息增强特征表达,提升伪标签质量和模型鲁棒性。结果表明,AVB-IS 在多种骨干网络中实现性能与效率的良好平衡,平均精度(AP)达到 49.35%,实例分割能力优于当前主流模型;同时,通过调节 KL 散度权重,模型展现出更快且更稳定的收敛特性。基于 AVB-IS 模型半监督学习(AVB-IS-SSL)框架下,仅使用 50% 的标注数据可获得接近全监督的效果,显著降低了对标注数据的依赖。本研究为玉米雄穗智能检测提供了可靠技术方案。

关键词:实例分割;半监督学习;贝叶斯;伪标签;玉米雄穗

中图分类号:TP391.41　　**文献标识码**:A　　**文章编号**:1000-1298(2026)12-0233-09　　**OSID**:

Instance Segmentation of Corn Tassels Based on Adaptive Variational Bayesian Semi-supervised Learning

WANG Qirui[1,2]　WANG Long[2]　SONG Shifang[3]　ZHOU Enquan[2,3]　MAO Hanping[3]　LIU Yang[4]

(1. *School of Mechanical Engineering*, *Changzhou Institute of Technology*, *Changzhou* 213032, *China*
2. *Haian Institute of Intelligent Equipment*, *Shanghai Jiao Tong University*, *Haian* 226600, *China*
3. *Key Laboratory of Modern Agriculture Equipment and Technology*, *Ministry of Education*, *Jiangsu University*, *Zhenjiang* 212013, *China*
4. *Mechanical Equipment Research Insistute*, *Xinjiang Academy of Agricultural and Reclamation Science*, *Shihezi* 832000, *China*)

Abstract: Aiming to address the problems of time-consuming and labor-intensive annotation in corn tassel instance segmentation, susceptibility to subjective factors, and high prediction uncertainty caused by complex field scenes, an adaptive variational-inference Bayesian instance segmentation model (AVB-IS) was proposed. AVB-IS integrated Bayesian probabilistic analysis with deep learning, learned the feature distribution of corn tassels through variational inference, and optimized the latent space with adaptive priors, enabling the model to predict uncertainty in a single forward pass. AVB-IS leveraged uncertainty information to enhance feature representation, improving pseudo-label quality and model robustness. Experimental results demonstrated that AVB-IS achieved an effective balance between performance and efficiency across diverse backbone networks, attaining an average precision (AP) of 49.35% and delivering instance segmentation performance that surpassed mainstream models. Moreover, adjusting the KL divergence weight enabled faster and more stable convergence. Under the framework of semi-supervised learning with the AVB-IS model (AVB-IS-SSL), the model achieved performance comparable to full supervision using only 50% of the annotated data, substantially reducing dependence on manual labeling. The research result can offer a robust technical solution for intelligent maize tassel detection and hold significant value for advancing agricultural intelligence.

Key words: instance segmentation; semi-supervised learning; Bayesian; pseudo-label; corn tassel

收稿日期:2026-01-10　修回日期:2026-01-26

基金项目:兵团重点领域科技攻关计划项目(2023AB015)

作者简介:王奇瑞(1983—),男,讲师,海安上海交通大学智能装备研究院博士后,主要从事智能农业装备技术研究,E-mail:wangqirui83@126.com

通信作者:刘洋(1978—),男,研究员,博士,主要从事玉米制种机械设计与应用研究,E-mail:Lyshz@163.com

0 引言

玉米去雄是杂交制种过程中的关键环节,及时、准确地去除雄穗,对于保障种子纯度和产量具有重要意义[1-3]。基于计算机视觉的雄穗识别与分割是实现智能化去雄的核心技术[3-4],但其依赖于大规模、高质量像素级标注数据,制作过程不仅耗时耗力,还易受人工主观影响[5]。

为降低对标注数据的依赖,半监督学习方法被引入识别任务中。早期的自训练半监督学习方法简单、直观,但易产生误差累积。Mean Teacher 针对自训练方法的缺陷引入正则化思想,为后续方法奠定了基础[6]。当前主流半监督学习方法主要基于一致性正则化和伪标签自训练机制[7-8],通过对无标签图像施加强、弱 2 种增强,并根据置信度预测生成伪标签[9]。然而,在目标形态多变、遮挡、特征粘连以及光照复杂的农业应用场景下[10-12],分割模型会对局部像素产生置信度较高但实际错误的预测。由于无法辨识此类认知模糊区域[13-14],学习过程中会引入错误的监督信号。这些误差在训练过程中不断被强化与传播[15],最终导致模型性能退化。概率深度学习方法为量化预测不确定性提供了可行思路[16-17]。蔡雨霖等[18]提出了一种半监督分割网络 UANP-MT(Uncertainty aware and network perturbed mean teacher),通过多次前向传播估计不确定性,用于提升目标分割与半监督学习的鲁棒性[19],但其计算量较大。高斯过程[20]理论完备,也因计算复杂而难以扩展至大规模视觉任务。为平衡效率与性能,WANG 等[21]提出一种基于神经过程的分割模型,通过恒定时间推理降低了计算成本,但其变分推理依赖预设先验条件,在复杂农业图像中应用受限。综上所述,现有不确定性建模方法难以在计算效率与农业场景的适应性之间取得良好平衡。

针对计算效率与适应性平衡难题,本文提出一种自适应变分推理贝叶斯实例分割模型。在统一的贝叶斯推理框架下,将高斯过程集成于神经网络,使不确定性估计从依赖多次前向传播的蒙特卡罗采样,简化为单次前向传播。通过变分推理将图像特征映射为潜在分布,以使模型能够从训练数据中隐式学习特征分布,并利用自适应机制将其转换为结构化的先验分布。以期提升模型计算效率,并通过量化预测不确定性动态增强特征,提升伪标签质量,降低错误预测在训练过程中误差累积。

1 材料与方法

1.1 数据集制作

数据来自新疆昌吉州玛纳斯县凉州户镇的玉米制种基地,采集对象为抽雄期玉米雄穗,拍摄设备为尼康 S3400D 型相机。所有图像均使用 Labelme 工具进行人工多边形标注,并转换为 COCO 数据格式。数据集中包含 300 组标注样本,按比例 14∶3∶3划分为训练集、验证集和测试集。如图 1 所示,数据集构建过程面临多重挑战:玉米雄穗结构复杂,在自然环境下常受叶片遮挡、特征粘连及光照不均等因素影响,使精准标注的难度及标注的时间成本增加;不同标注人员对雄穗结构的理解和操作习惯存在差异,易造成标注标准不一致,从而影响模型对目标特征的稳健学习。

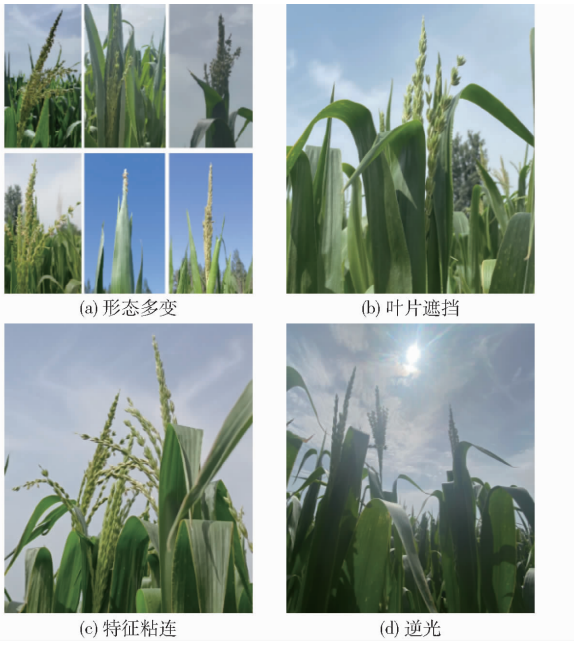

图 1 试验样本示例

Fig. 1 Representative examples of test samples

1.2 自适应变分推理贝叶斯实例分割理论建模

1.2.1 贝叶斯分割概率模型

传统实例分割模型建立了一个确定性映射,计算式为

$$\boldsymbol{y}=f(\boldsymbol{x}) \tag{1}$$

贝叶斯分割模型将分割问题转换为概率生成过程,计算式为

$$\boldsymbol{y}\sim p(\boldsymbol{y}|\boldsymbol{x})=E_{z\sim p(\boldsymbol{y}|\boldsymbol{x})}[p(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})]=\int p(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})p(\boldsymbol{z}|\boldsymbol{x})\,\mathrm{d}z \tag{2}$$

式中 $p(\boldsymbol{z}|\boldsymbol{x})$——输入特征 $\boldsymbol{x}$ 时,潜在特征变量 $\boldsymbol{z}$ 的条件先验分布

$p(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})$——似然函数,在已知潜在变量 $\boldsymbol{z}$ 和输入 $\boldsymbol{x}$ 的情况下,生成分割掩码 $\boldsymbol{y}$ 的概率

$p(\boldsymbol{y}|\boldsymbol{x})$——边缘似然

通过学习一个概率分布,模型能更有效地提取特征,还能量化预测中的不确定性,从而提升对分布

外样本的鲁棒性。在贝叶斯分割概率建模中，对于输入特征 $\boldsymbol{x}$ 和分割掩码 $\boldsymbol{y}$，引入中间潜在变量 $\boldsymbol{z}$ 刻画分割中的不确定性。$\boldsymbol{z}$ 为低维潜在特征增强向量，具有语义意义，其编码了边界模糊、遮挡、尺度变化等固有不确定性因素。引入 $\boldsymbol{z}$ 后，模型不再局限于单点估计，而能刻画数据分布的内在规律，从而在复杂场景中表现出更强的适应性与鲁棒性。

1.2.2　变分推理近似求解

当变量 $\boldsymbol{z}$ 为高维连续变量时，边缘似然 $p(\boldsymbol{y}|\boldsymbol{x})$ 中的积分难以直接求解，因此需采用变分推理进行近似求解。为此，引入由参数 ϕ 控制的变分后验分布 $q_\phi(\boldsymbol{z}|\boldsymbol{x})$，用以逼近真实后验 $p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{y})$，变分下界为

$$\ln p(\boldsymbol{y}|\boldsymbol{x}) \geqslant E_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\ln p_\theta(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})] - D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x}) \parallel p(\boldsymbol{z})) \tag{3}$$

式中　$E_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\ln p_\theta(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})]$——重构损失，衡量在从变分后验 $q_\phi(\boldsymbol{z}|\boldsymbol{x})$——采样得到 $\boldsymbol{z}$ 的情况下，模型重构目标 $\boldsymbol{y}$ 的能力

$D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x}) \parallel p(\boldsymbol{z}))$——KL 损失，度量变分后验 $q_\phi(\boldsymbol{z}|\boldsymbol{x})$ 与先验 $p(\boldsymbol{z})$ 之间的差异

在训练阶段，最大化变分下界同时优化生成模型参数 θ 与推理网络参数 ϕ，计算式为

$$\theta,\phi = \arg\max_{\theta,\phi}\{E_{q_\phi(\boldsymbol{z}|\boldsymbol{x})}[\ln p_\theta(\boldsymbol{y}|\boldsymbol{z},\boldsymbol{x})] - D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{x}) \parallel p(\boldsymbol{z}))\} \tag{4}$$

1.2.3　自适应先验分布

传统变分自编码器通常采用固定的标准高斯先验 $p(\boldsymbol{z}) = N(0, I)$，但这种简单假设表达能力有限，难以与真实的复杂数据结构相匹配，也无法利用训练过程中学到的后验信息。自适应先验机制在保持贝叶斯理论框架优势的同时，能够更好地利用具体任务的数据，充分学习训练过程获得的分布信息，从而提升模型表示能力与生成性能。自适应先验的机制为

$$p(\boldsymbol{z})^{t+1} = \alpha p(\boldsymbol{z})^{t} + (1-\alpha)E[q_\phi(\boldsymbol{z}|\boldsymbol{x})] \tag{5}$$

式中　$p(\boldsymbol{z})^{t}$——第 t 次迭代时的先验分布

$E[q_\phi(\boldsymbol{z}|\boldsymbol{x})]$——当前数据后验分布期望

α——动量参数

1.3　AVB－IS 模型

AVB－IS 模型由骨干网络、区域提议网络及贝叶斯分割头组成。骨干网络作为特征提取器从输入图像中提取特征；区域提议网络在骨干网络生成的特征图上产生候选区域；贝叶斯分割头将候选区域的特征图经多尺度特征卷积、贝叶斯编码、重参数化、注意力机制、多特征融合和最终预测卷积输出分割结果。如图 2 所示，自适应贝叶斯分割头以感兴趣区域(ROI)对齐后的特征图为输入，包括训练与推理 2 种模式：训练模式在训练阶段依托真实标签或伪标签计算损失，推理模式则用于生成最终预测结果。

1.3.1　训练模式

如图 2 所示，在训练阶段，给定一批标注数据

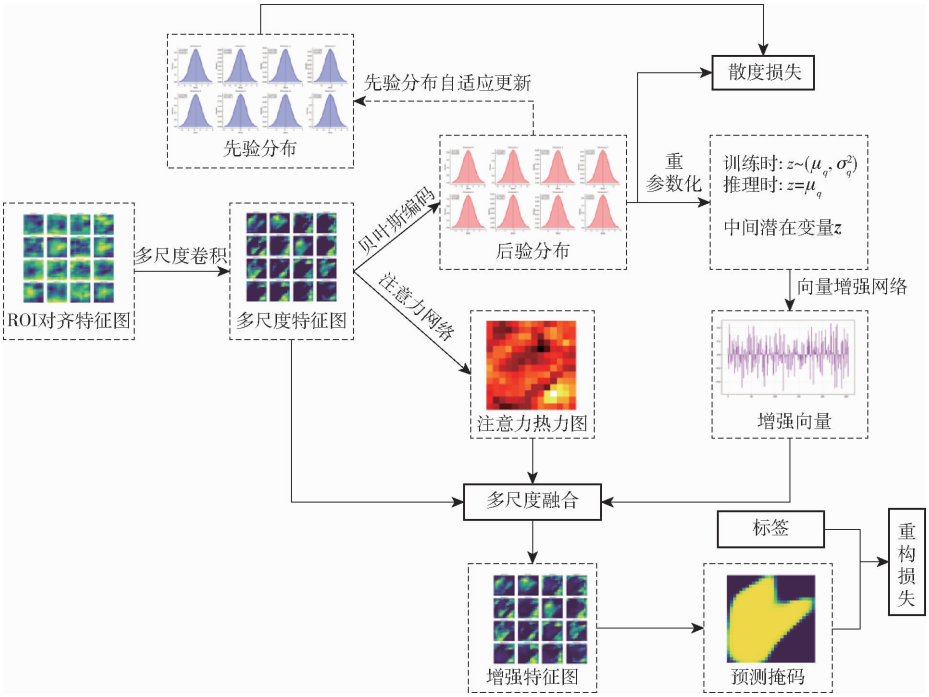

图 2　自适应变分推理贝叶斯分割的工作流程与原理

Fig. 2　Workflow and principles of adaptive variational Bayesian inference-based segmentation

后,自适应贝叶斯分割头通过多尺度卷积提取输入特征图的层次化表示。随后分割头并行处理 2 条分支:贝叶斯增强分支通过变分推理生成带有不确定性估计的增强特征;确定性分支通过标准卷积操作维持稳定的特征表示。2 条分支的输出通过注意力机制进行特征融合,解码器在此基础上计算重构损失和 KL 散度损失。

贝叶斯增强路径,输入特征通过多尺度卷积进行处理,得到多尺度特征图。多尺度特征图通过全局平均池化进行空间压缩,再由 1 个编码器网络处理,输出后验分布(μ,σ^2)。模型中初始化了一组可学习的先验分布参数,这些参数在训练过程中根据后验分布进行动量更新:随着迭代次数增加,更新动量从 0.9 逐渐提升至 0.999,确保先验分布逐步适应数据分布。通过重参数化,训练时从后验分布(μ,σ^2)中采样潜在向量,推理时直接使用均值向量。这些潜在向量通过增强器网络生成特征增强向量,其增强强度由 KL 权重控制。

确定性路径输入多尺度特征图,通过卷积注意力机制生成空间注意力图。将贝叶斯增强向量与注意力图相乘后加权调制主特征;随后通过可学习的残差参数将不同尺度的特征经上采样后以残差方式融合到增强特征中,形成多尺度增强的特征表示。

融合后的特征通过转置卷积进行上采样,最终由预测层生成逐像素的掩码。不确定性通过 KL 散度损失隐式学习,计算后验分布与自适应更新后的先验分布之间的散度。为节省存储空间,训练后仅保存先验分布参数用于推理,不存储完整的后验分布。

1.3.2 推理模式

推理模式下,输入的测试数据,通过预训练的多尺度卷积网络提取多尺度特征图;经过全局平均池化压缩,使用编码器网络获取潜在空间参数,并直接采用后验均值作为潜在向量(跳过重参数化采样),同时结合训练阶段学习到的先验分布参数,通过增强器网络生成特征增强向量;将贝叶斯增强特征与多尺度卷积特征通过注意力机制进行融合整合多尺度特征,由解码器对融合后的特征进行上采样和逐像素预测,整个推理过程通过消除随机采样和固定先验参数,实现了高效、稳定的预测性能。

1.4 基于 AVB-IS 模型的自训练学习

基于少量已标注数据训练初始的 AVB-IS 模型。使用该模型对大量未标注数据进行预测,并将其中预测置信度超过预设阈值的样本视为可靠性较高,为其分配伪标签。将伪标签样本与原有标注数据合并,构成扩增后的训练集,用于对 AVB-IS 模型进行重新训练。通过迭代模型不断提升其精度与鲁棒性。

1.5 超参数设置

模型关键参数设置为:贝叶斯推理层的潜在维度设为 8,基础学习率为 0.005。为平衡稳定特征与增强特征表示,KL 散度权重采用渐进式策略,从 0 开始,在训练的前 50% 阶段线性增加至目标值。此外,将检测置信度阈值设为 0.95,并对预测掩模施加最小 800 像素的面积过滤。

1.6 评价指标

模型评估以掩码平均精度 AP 为核心指标,重点关注 AP50 与 AP75 以综合衡量模型在不同 IoU 阈值下定位准确性与掩码质量;引入浮点运算量、参数量以及帧率等效率指标,构成兼顾性能与复杂度的综合评价。

1.7 试验平台

试验平台搭载 Intel(R) Core(TM) i7-14700F 处理器、NVIDIA GeForce RTX 5060 Ti 显卡、64 GB 系统内存及 932 GB SSD 固态硬盘。操作系统为 64 位 Windows,主要开发环境为 Python 3.9、PyTorch 2.7.1 及 CUDA 12.8。

2 结果与讨论

2.1 骨干网络对 AVB-IS 模型性能的影响

为研究骨干网络结构对 AVB-IS 分割性能的影响,在测试集上评估了 ResNet-101、ResNet-50、ResNet-34 及 EfficientNetV2 的分割表现,结果如表 1 所示。由表 1 可知,整体上 ResNet 系列呈现模型规模与性能正相关趋势。ResNet-101 + DC5 在参数量与浮点运算量最高时,取得了最优 AP。值得注意的是,ResNet-50 + FPN 在显著降低参数量(4.46×10^7)和浮点运算量(2.15×10^{11})的情况下,仍保持了与 ResNet-101 + DC5 相近的 AP 表现(49.35%),在提高计算效率下仍能保持较高分割性能。然而,当网络深度与容量进一步减小至 ResNet-34 + FPN 时,分割性能出现明显下滑,AP 降至 36.87%,尤其在对定位精度要求更高的 AP75 上仅为 22.19%,表明较浅网络在高精度语义特征提取方面存在明显不足。另一方面,尽管 Swin-S、ConvNeXt-S 和 EfficientNetV2 的参数量与 ResNet-50 相近,计算量也相差不大,但其各项性能指标均显著偏低,说明其网络结构不适于当前分割任务。MobileNetV3-S 虽然有极低的参数量和计算复杂度但其 AP 最低。综合考虑模型分割性能与计算效率,ResNet-50 + FPN 在保持较高精度的同时具备更低的计算复杂度,因此选用其作为骨干网络更为合理。

表 1　不同骨干网络 AVB－IS 模型性能测试结果

Tab. 1　Performance test results of AVB－IS model with different backbone networks

模型	参数量	浮点运算量	AP50/%	AP75/%	AP/%
ResNet－101＋DC5	6.38×10^{7}	3.60×10^{11}	94.45	45.46	49.41
ResNet－101＋FPN	6.37×10^{7}	2.90×10^{11}	92.31	45.32	49.00
ResNet－50＋DC5	4.47×10^{7}	2.50×10^{11}	95.24	44.37	48.14
ResNet－50＋FPN	4.46×10^{7}	2.15×10^{11}	92.76	45.46	49.35
ResNet－34＋FPN	4.40×10^{7}	4.30×10^{10}	88.81	22.19	36.87
Swin－S	5.98×10^{7}	2.05×10^{11}	57.24	7.30	21.87
ConvNeXt－S	3.82×10^{7}	2.42×10^{11}	76.17	8.70	26.70
MobileNetV3－S	1.80×10^{7}	1.70×10^{10}	58.89	6.54	17.07
EfficientNetV2	4.75×10^{7}	1.79×10^{11}	51.52	16.46	21.34

2.2　KL 权重调控下模型收敛特性与性能分析

图 3a 为不同 KL 权重下模型在实例分割任务上的性能表现。在宽松定位要求下，AP50 在所有 KL 权重条件下均保持高水平且波动较小，说明 AP50 对 KL 正则化强度的变化不敏感。相比之下，AP75 随 KL 权重变化波动较大，表明 KL 正则化对高精度定位性能具有显著影响。AP 为 47.43% ~ 49.55%，并在 KL 权重为 0.1 时达到峰值。因此，将 KL 权重设为 0.1 可维持较高 AP50、AP75，优化整体 AP 之间取得较好平衡。

在变分推理框架中，KL 散度损失用于度量近似后验分布与先验分布之间的差异。图 3b 为模型训练中 KL 散度损失随迭代次数的变化趋势，整体可划分为 3 个阶段：初始阶段（0 ~ 100 次迭代），KL 散度损失从 0.18 开始快速下降，意味着近似后验与分布之间的差异迅速减小，；中期阶段（100 ~ 1 000 次迭代），下降速度逐渐减缓，模型逼近最优解；后期阶段（1 000 次迭代之后），KL 散度损失稳定在 0.005 以下，仅因优化随机性产生轻微波动，显示模型已基本收敛。

图 3c 为掩码损失同样呈现 3 阶段收敛特征：初始阶段（0 ~ 1 000 次迭代），模型尚未具备有效重构能力，损失维持在较高水平；中期阶段（1 000 ~ 2 500 次迭代），损失以指数衰减形式缓慢下降，此时模型已学习到主体结构，优化重点转向细节，波动主要源于批次采样与参数更新的随机性；后期阶段（2 500 次迭代之后），损失稳定在 0.132 8 ~ 0.203 0，表明模型已收敛，后续波动主要由批处理噪声和优化随机性引起。

图 3d 为验证集上平均精度（AP）随迭代次数的

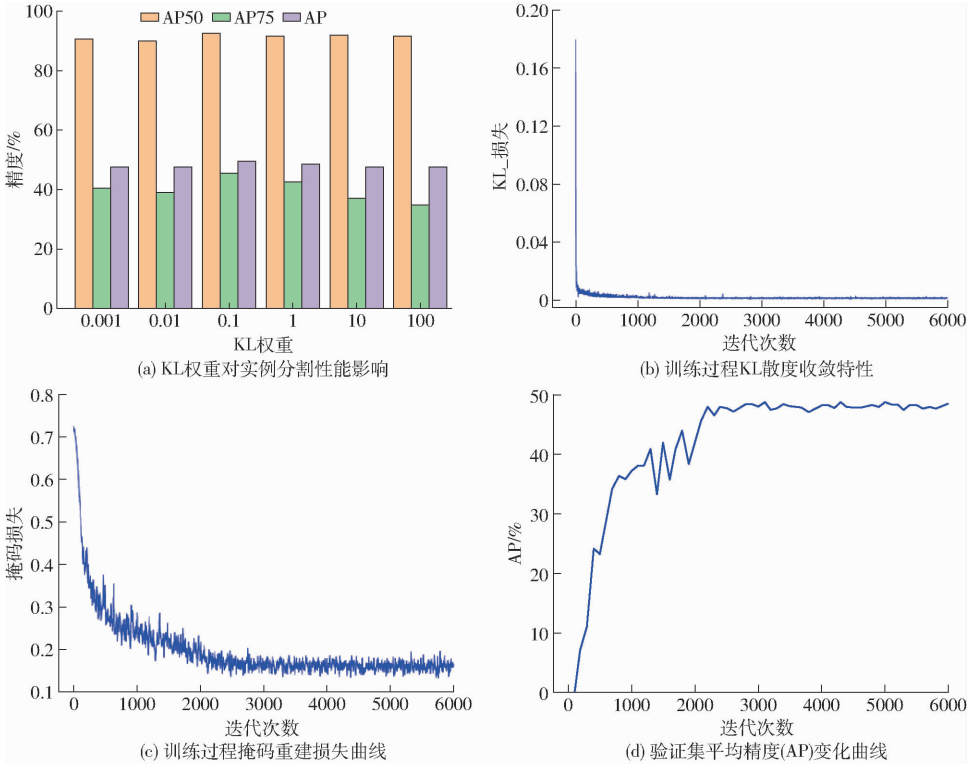

图 3　KL 正则化与模型收敛分析

Fig. 3　Analysis of KL regularization and model convergence

演变,其趋势与掩码损失高度一致:迭代 0 ~ 1 000 次内 AP 快速上升;迭代 1 000 ~ 2 500 次增速放缓且波动增大,反映模型正在学习更复杂的判别特征并增强场景适应性;迭代 2 500 次后 AP 基本稳定在 48% ~ 49%,标志模型训练完成。KL 散度损失在训练早期快速收敛并持续保持低位,这种快速而稳定的收敛有助于抑制过拟合,并促进潜在空间形成平滑、连续且结构清晰的流形。在此更具几何结构性与噪声鲁棒性的潜空间中,掩码重建任务能更有效地解耦语义信息、降低重建难度,进而推动模型整体性能提升。KL 权重通过调节正则化强度影响后验分布的收敛特性,进而作用于掩码重建过程与模型整体性能。合适的 KL 权重有助于在训练稳定性、重建质量与检测精度之间实现最优平衡。

2.3 AVB - IS 与主流实例分割模型性能

表 2 为采用 ResNet - 50 作为骨干网络时,对比了 AVB - IS 模型与多种主流实例分割模型(Mask R - CNN[22]、MS R - CNN[23]、Cascade R - CNN[24]、HTC[25]、SOLO[26] 和 SOLOv2[27])在迭代 3 000 次与 6 000 次下的性能,并在输入分辨率为 800 像素 × 1 333 像素的条件下测试了各模型的推理速度。整体来看,AVB - IS 在精度、效率与收敛速度均有一定优势。在精度方面,AVB - IS 仅迭代 3 000 次 AP 达到 48.51%,明显高于其他模型;迭代 6 000 次后 AP 提升至 49.35%,同时在 AP50(92.74%)与 AP75(45.46%)也领先,展现出更强的特征表达与分割性能。在效率方面,尽管单阶段模型如 SOLO 系列在参数量与浮点运算量更轻量,且具有较高推理速度,但其整体性能仍劣于 AVB - IS 及双阶段模型。Cascade R - CNN 与 HTC 虽参数量和浮点运算量更大,却未带来相应的性能提升。相比之下,AVB - IS 的参数量约为 4.450×10^7,浮点运算量为 2.15×10^{11},显著低于 Cascade R - CNN 与 HTC,说明其优势主要源于高效的特征融合与重参数化设计,而非依赖参数量堆叠。在收敛速度方面,AVB - IS 仅需迭代 3 000 次就在关键指标上全面领先,体现了更快的学习效率与更稳定的优化过程,有助于大幅降低训练时间与计算成本。

表 2 AVB - IS 与主流模型性能测试结果

Tab. 2 Performance test results of AVB - IS and mainstream models

模型	迭代次数	AP50/%	AP75/%	AP/%	参数量	帧率/(f·s^{-1})	浮点运算量
Mask R - CNN	3 000	84.09	19.79	37.16	4.42×10^7	13.93	2.07×10^{11}
	6 000	87.06	28.87	44.85			
MS R - CNN	3 000	88.00	21.30	36.50	4.50×10^7	12.85	2.25×10^{11}
	6 000	89.70	34.10	43.50			
Cascade R - CNN	3 000	86.20	23.70	37.30	5.15×10^7	12.61	2.95×10^{11}
	6 000	88.60	32.80	43.40			
HTC	3 000	90.90	20.80	37.60	5.71×10^7	12.26	3.35×10^{11}
	6 000	88.10	36.20	43.10			
SOLO	3 000	53.60	9.60	19.10	3.30×10^7	14.12	2.00×10^{11}
	6 000	65.20	19.60	27.80			
SOLOv2	3 000	90.20	20.40	37.60	3.10×10^7	14.77	1.90×10^{11}
	6 000	88.10	26.20	43.10			
AVB - IS	3 000	93.02	42.87	48.51	4.45×10^7	13.33	2.15×10^{11}
	6 000	92.74	45.46	49.35			

图 4 展示了表 2 各分割模型在复杂田间背景下,对玉米雄穗分割效果的可视化。测试图像(图 4a)包含植株密集、目标多尺度分布及叶片显著遮挡等特征,构成典型农业场景下的示例分割挑战。由图 4b ~ 4h 可知,AVB - IS 模型表现最优,成功分割出 8 个雄穗实例,其掩码边界贴合度高,尤其在处理部分遮挡与小尺度目标时展现出细节保留能力;相比之下,SOLO 模型仅分割出 5 个实例,存在明显漏检,而 SOLOv2、MS R - CNN 及 HTC 等主流方法均分割出 6 个实例,Mask R - CNN 和 Cascade R - CNN 模型分割出 7 个实例,虽能识别主要目标,但在密集区域易出现边界模糊或实例粘连现象。可视化结果表明,AVB - IS 在复杂自然背景下实现了更鲁棒且精细的实例分割性能,显著优于现有通用模型。

2.4 不同标注比例下自训练与全监督方法性能测试

由表 3 可知,在标注数据比例为 10% ~ 50% 条件下,自训练半监督学习方法的 AP、AP50 和 AP75 均稳定优于全监督学习方法,说明未标注数据被有效利用,从而提升了模型性能。当标注数据仅占 10% 时,自训练方法在 AP 上较全监督方法提升

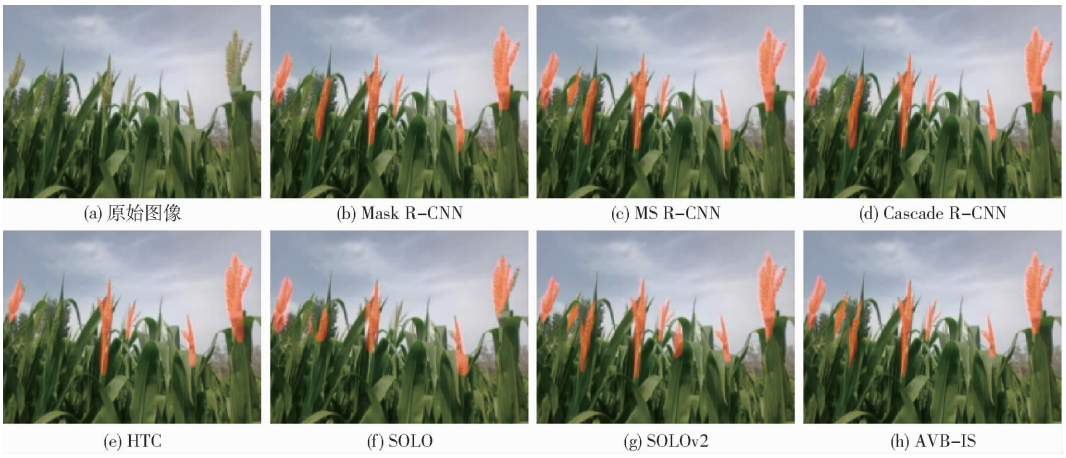

图 4　复杂背景下各模型实例分割效果可视化

Fig. 4　Visual of instance segmentation performance across models under complex backgrounds

1.82 个百分点，在 AP50 上提升 4.72 个百分点，表明模型能够在少量标注与大量未标注数据的协同下学习到更稳健的目标特征。随着标注比例提高，自训练方法的优势持续保持，标注数据提升到 30% 时尤为明显：AP 提升 3.83 个百分点，AP75 提升 7.56 个百分点，说明半监督机制在提升高精度定位（IoU ≥ 75%）方面效果显著。当标注比例达到 50% 时，自训练方法 AP 为 49.08%，AP50 为 93.79%，AP75 为 46.22%，已接近使用 100% 标注数据的全监督学习水平。

表 3　不同标注比例下全监督与自训练方法性能测试结果

Tab. 3　Performance test results of fully-supervised and self-training methods at different annotation ratios

%

数据比例/%	模式	AP	AP50	AP75
10	全监督	39.33	86.36	27.62
	自训练	41.15	91.08	28.15
20	全监督	43.08	90.54	30.84
	自训练	44.26	90.96	35.96
30	全监督	42.29	88.33	33.55
	自训练	46.12	91.29	41.11
40	全监督	44.68	92.44	36.54
	自训练	47.68	92.55	42.43
50	全监督	45.39	90.99	37.86
	自训练	49.08	93.79	46.22

图 5 每幅子图左侧为人工标注标签，右侧为使用 30% 标注数据训练的 AVB－IS 模型生成的伪标签，从单目标标注、复杂边界标注、多目标标注以及部分遮挡目标标注等 4 个典型任务出发，对 AVB－IS 模型的标注性能进行对比分析。由图 5a 可知，AVB－IS 生成的伪标签（右）在目标位置与尺度上均与人工标注高度一致，仅顶部存在轻微偏移，表明模型在规则形状的单目标分割任务中具有可靠定位能力。由图 5b 可知，AVB－IS 生成的伪标签（右）

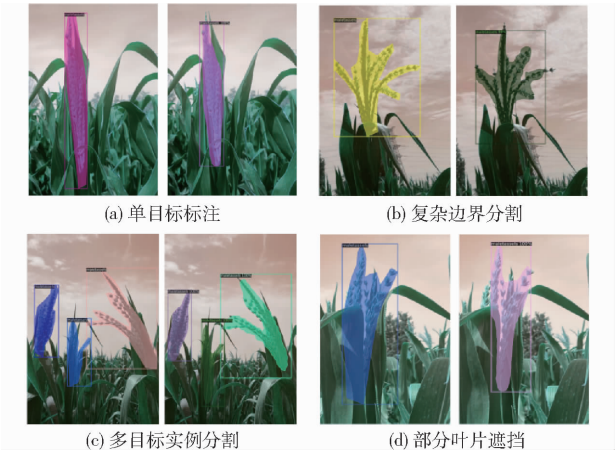

图 5　人工标注与 AVB－IS 生成伪标签

Fig. 5　Manual annotation and pseudo-labels generated by AVB－IS

对细微分支结构的捕捉与人工标注存在差异，尤其在轮廓边缘形状和花序末端细节部分，导致轮廓信息不够完整，但基本标注出了目标主体。由图 5c 可知，AVB－IS 生成的伪标签（右）能正确区分不同雄穗实例，并分配独立实例标签，各目标之间无交叉或混淆，展现了良好的实例区分性能。由图 5d 可知，AVB－IS 模型（右）未能完整重建被遮挡部分的几何形状，导致目标掩码面积较人工标注偏小，说明在遮挡恢复方面仍有改进空间。由图 5 可知，AVB－IS 在农业图像自动标注任务中展现出潜力：在单目标、多目标结构相对简单场景下，其表现已接近人工标注水平；面对复杂边界与局部遮挡等挑战性场景时，虽存在细节丢失与形状估计偏差，仍能实现实例的有效标注。

2.5　AVB－IS－SSL 在半监督实例分割任务上性能测试

表 4 为在 Cityscapes 数据集上仅使用 10% 标注数据进行半监督学习时，多种半监督实例分割方法的性能表现。由表 4 可知，AVB－IS－SSL 的 AP50

为 57.38%，AP 为 31.58%，整体性能优于 Self-training、Unbiased Teacher、PAIS、Teacher 和 Guided Distillation 等主流方法。即使在计算开销显著低于 CAST 的情况下，AVB－IS－SSL 的分割精度依然优于 CAST。这一性能提升主要得益于其核心模型 AVB－IS 所引入的自适应变分推理贝叶斯框架。通过对预测不确定性建模，AVB－IS 能够有效识别并抑制低质量伪标签带来的噪声干扰，从而保持稳定且鲁棒的特征学习能力。在半监督训练过程中，这种不确定性感知机制显著增强了伪标签的可靠性。

表 4 AVB－IS－SSL 与主流方法性能测试结果

Tab. 4 Performance test results of AVB－IS－SSL and mainstream methods

模型	参数量	浮点运算量	AP50/%	AP/%
Self-training[28]	6.60×10^7	3.70×10^{10}	54.90	29.70
Unbiased Teacher[29]	4.42×10^7	2.07×10^{11}	54.90	29.80
PAIS[30]	4.42×10^7	2.07×10^{11}	44.90	22.90
Guided Distillation[31]	5.50×10^7	3.50×10^{11}	52.90	30.80
CAST (teacher adaptation)[32]	5.68×10^8	8.80×10^{11}	56.60	30.50
AVB－IS－SSL	4.45×10^7	2.15×10^{11}	57.38	31.58

3 结论

(1)提出了一种基于自适应变分贝叶斯的实例分割模型 AVB－IS，将贝叶斯概率框架融入深度网络，可在单次前向传播中实现预测不确定性估计，为复杂农业场景下实例分割提供了新的解决思路。模型通过变分推断隐式学习数据特征分布，并结合自适应先验机制动态优化潜在空间，从而实现不确定性量化并强化特征表达，最终提升伪标签生成质量。结构上以 ResNet－50 与 FPN 作为骨干网络，在性能与效率之间取得了良好平衡。

(2)AVB－IS 在玉米雄穗实例分割任务中的平均精度(AP)达到 49.35%，显著优于 Mask R－CNN、Cascade R－CNN 等主流模型。将 KL 散度权重设为 0.1 可使模型快速且稳定收敛，有效平衡特征重建与分布正则化约束。结合自训练机制后，仅依赖 30% ~50% 的标注数据即可获得接近全监督学习的性能，大幅降低了对人工标注的需求。

(3)本文模型在保持高计算效率的同时，显著提升了复杂田间场景中的实例分割精度与鲁棒性，为玉米雄穗智能化去雄提供了有力的技术支持，对推动农业自动化进程、降低人工成本具有重要意义。

参 考 文 献

[1] 王国新，韩科立，邓志刚，等. 自走式玉米去雄机设计与试验[J]. 农业工程，2020，10(2)：67－71.
WANG Guoxin，HAN Keli，DENG Zhigang，et al. Design and test of self-propelled corn emasculation machine[J]. Agricultural Engineering，2020，10(2)：67－71.(in Chinese)

[2] ZHANG R R，YANG J X，CHEN L P，et al. Review of the detasseling techniques for maize (*Zea mays* L.) hybrid seed production[J]. International Journal of Agricultural and Biological Engineering，2024，17(3)：1－11.

[3] 李金瑞. 基于深度学习的制种玉米雄穗检测方法研究[D]. 杨凌：西北农林科技大学，2025.
LI Jinrui. Research on detection method of seed corn tassel based on deep learning[D]. Yangling：Northwest A&F University，2025.(in Chinese)

[4] 杨佳璇，陈立平，张瑞瑞，等. 基于机器视觉的制种玉米遗漏雄穗识别检测[J]. 农机化研究，2026，48(1)：182－189，198.
YANG Jiaxuan，CHEN Liping，ZHANG Ruirui，et al. Identification and detection of omitted tassels in maize for seed based on machine vision[J]. Journal of Agricultural Mechanization Research，2026，48(1)：182－189，198.(in Chinese)

[5] ALZUBAIDI L，BAI J S，AL-SABAAWI A，et al. A survey on deep learning tools dealing with data scarcity：definitions，challenges，solutions，tips，and applications[J]. Journal of Big Data，2023，10(1)：46.

[6] 陈龙伟，周小成，李传昕，等. 基于 UNet－ResNet14* 半监督学习的无人机影像森林树种分类[J]. 农业工程学报，2024，40(1)：217－226.
CHEN Longwei，ZHOU Xiaocheng，LI Chuanxin，et al. Classification of tree species based on UNet－ResNet14* semi-supervised learning using UAV images[J]. Transactions of the CSAE，2024，40(1)：217－226.(in Chinese)

[7] 刘涛，韩伊曼，杨奉源，等. 基于无人机遥感影像土壤背景剔除的玉米叶部病害等级识别模型[J]. 农业机械学报，2025，56(11)：408－417.
LIU Tao，HAN Yiman，YANG Fengyuan，et al. Classification of maize leaf disease level information from UAV remote sensing images with soil background removal[J]. Transactions of the Chinese Society for Agricultural Machinery，2025，56(11)：408－417.(in Chinese)

[8] SOHN K，BERTHELOT D，LI C L，et al. FixMatch：simplifying semi-supervised learning with consistency and confidence[J]. Advances in Neural Information Processing Systems，2020，33：596－608.

[9] 丁永军，杨文涛，赵一龙. 基于一致性半监督学习的苹果叶片病斑分割模型研究[J]. 农业机械学报，2024，55(12)：314－321.
DING Yongjun，YANG Wentao，ZHAO Yilong. Apple leaf spot segmentation model based on consistency semi-supervised learning[J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(12)：314－321.(in Chinese)

[10] WANG J. Exploring probabilistic models for semi-supervised learning[D]. Oxford: University of Oxford, 2023.
[11] 赵亚楠, 邓寒冰, 刘婷, 等. 基于弱监督学习的玉米苗期植株图像实例分割方法[J]. 农业工程学报, 2022, 38(19): 143 - 152.
ZHAO Ya'nan, DENG Hanbing, LIU Ting, et al. Instance segmentation method of seedling maize plant images based on weak supervised learning[J]. Transactions of the CSAE, 2022, 38(19): 143 - 152. (in Chinese)
[12] 翟雪东, 韩文霆, 马伟童, 等. 基于改进 AdvSemiSeg 的半监督遥感影像作物制图方法[J]. 农业机械学报, 2024, 55(8): 196 - 204.
ZHAI Xuedong, HAN Wenting, MA Weitong, et al. Semi-supervised network for remote sensing crop mapping based on improved AdvSemiSeg[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(8): 196 - 204. (in Chinese)
[13] 车红岩,潘耀忠,夏兴生,等.基于集成和置信学习的半监督作物遥感分类方法[J].农业机械学报,2025,56(11):418 - 432.
CHE Hongyan, PAN Yaozhong, XIA Xingsheng, et al. Semi-supervised crop remote sensing classification based on ensemble and confident learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 418 - 432. (in Chinese)
[14] KAMILARIS A, PRENAFETA-BOLDU F X. Deep learning in agriculture: a survey[J]. Computers and Electronics in Agriculture, 2017, 147: 70 - 90.
[15] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. Journal of Machine Learning Research, 2016, 17(59): 1 - 35.
[16] 李万鹏. 一种基于深度学习和概率图模型的本征分解方法的研究[D]. 杭州: 浙江大学, 2018.
LI Wanpeng. A study on intrinsic decomposition method based on deep learning and probabilistic graphical models[D]. Hangzhou: Zhejiang University, 2018. (in Chinese)
[17] KENDALL A, GAL Y. What uncertainties do we need in Bayesian deep learning for computer vision? [C]// Advances in Neural Information Processing Systems, 2017:5574 - 5584.
[18] 蔡雨霖,肖佳仪,余超然,等. 基于 UANP - MT 的半监督菜心杂草分割方法[J]. 农业工程学报,2023,39(11):183 - 191.
CAI Yulin, XIAO Jiayi, YU Chaoran, et al. UANP - MT based semi-supervised image segmentation method for identifying weeds in cabbage field[J]. Transactions of the CSAE, 2023, 39(11): 183 - 191. (in Chinese)
[19] ZHANG Y C, JIAO R S, LIAO Q C, et al. Uncertainty-guided mutual consistency learning for semi-supervised medical image segmentation[J]. Artificial Intelligence in Medicine, 2023, 138: 102476.
[20] 张伟荣, 陈学庚, 齐江涛, 等. 基于深度学习和高斯过程回归的玉米冠下视觉导航路径提取方法[J]. 农业机械学报, 2024, 55(7): 15 - 26
ZHANG Weirong, CHEN Xuegeng, QI Jiangtao, et al. Deep learning and Gaussian process regression based path extraction for visual navigation under canopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(7): 15 - 26. (in Chinese)
[21] WANG J, LUKASIEWICZ T, MASSICETI D, et al. NP - Match: when neural processes meet semi-supervised learning[C]// International Conference on Machine Learning, 2022: 22919 - 22934.
[22] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R - CNN[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386 - 397.
[23] HUANG Z, HUANG L, GONG Y, et al. Mask scoring R - CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 6402 - 6411.
[24] CAI Z, VASCONCELOS N. Cascade R - CNN: delving into high quality object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154 - 6162.
[25] CHEN K, PANG J, WANG J, et al. Hybrid task cascade for instance segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 4974 - 4983.
[26] WANG X, KONG T, SHEN C, et al. SOLO: segmenting objects by locations[C]// European Conference on Computer Vision, 2020: 649 - 665.
[27] WANG X, ZHANG R, KONG T, et al. SOLOv2: dynamic and fast instance segmentation[C]// Advances in Neural Information Processing Systems, 2020: 17721 - 17732..
[28] XIE Q, LUONG M T, HOVY E, et al. Self-training with noisy student improves ImageNet classification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10687 - 10698.
[29] LIU Y C, MA C Y, HE Z, et al. Unbiased teacher for semi-supervised object detection[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 8368 - 8377.
[30] HU J, CHEN C, CAO L, et al. Pseudo-label alignment for semi-supervised instance segmentation[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 16337 - 16347.
[31] BERRADA T, COUPRIE C, ALAHARI K, et al. Guided distillation for semi-supervised instance segmentation[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 475 - 483.
[32] TAGHAVI P, LIU T, LI R, et al. CAST: contrastive adaptation and distillation for semi-supervised instance segmentation [J]. arXiv Preprint, arXiv:2505.21904, 2025.